Stochastic rainfall interpolation and downscaling

L. G. Lanza, J. A. Ramírez, E. Todini

Hydrology and Earth System Sciences, 5(2), *139–143* (2001) © EGS

Hydrology & Earth System Sciences

Stochastic rainfall interpolation and downscaling

L.G. Lanza[1], J.A. Ramírez[2] and E. Todini[3]

[1]University of Genova, Dept. of Environmental Engineeering, 1 Montallegro, 16145 Genova, Italy. Email:luca@diam.unige.it

[2] Water Resources and Hydrologic Sciences Civil Engineering Dept., Colorado State University, Fort Collins, C0 80523-1272, USA. EmailL ramirez@engr.ColoState.Edu

[3] Dept of Earth and Geo-Environmental Sciences, University of Bologna, Via Zambone, 67 40127 Bologna, Italy. Email: todini@geomin.unibo.it.

Introduction

Practising hydrologists and researchers have always been faced with the frustrating experience of the lack of suitable tools for the direct measurement of areal precipitation – the essential driving input of all processes involved in the ground phase of the water cycle. Indeed, while the rainfall process is known to exhibit a high degree of variability both in space and time, reliable direct rainfall measurements can be obtained only at the very limited spatial scale where rain catching is practically possible (often referred to in hydrology as the point scale).

Based on the assumption that rain gauge measurements can reliably account for the "true point rainfall" after accounting for a number of possible errors (see e.g. Sevruk, 1982; WMO, 1994; Humphrey *et al.*, 1997 and La Barbera *et al.*, 2001 for hints on the error sources), areal rainfall estimates have been obtained traditionally through some suitable interpolation method and aggregation technique. These are based on the hypothesis that rainfall estimates at ungauged sites can be obtained as linear or non-linear combinations of the values measured at a number of instrumented locations using the appropriate weights. Most of such approaches can be traced back from the subjective isohyets methods (Linsley *et al*., 1949), to the more objective geometrical interpolators based on Thiessen polygons (Linsley *et al*., 1949, WMO, 1986), to the mathematical surface interpolators based on splines (Matheron, 1981), to the paradigms of geostatistics – the new branch of statistical theory introduced by Matheron (1971) in the field of mining engineering.

On the other hand, indirect estimates of areal rainfall based on the measurement of related ancillary variables have been provided since the late '60s by ground-based meteorological RADARs and remote sensing devices borne on satellite platforms, such as RADARs and other sensors, with varied associated degrees of reliability. All such indirect methods, whose significance and reliability for hydrological purposes are still to be determined, are often assumed to provide an indication of the coarse scale pattern of the rain field useful in many applications. Nevertheless, indirect methods must first be calibrated and validated using historical data and then, in operation, adjusted continuously against direct ground measurements (e.g. rain gauges).

Finally, physically-based numerical models of the atmosphere — though relying on various theoretical approximations — provide predictions of temporal accumulation values for areal rainfall over wide spatial scales. The output of such models has little to do with actual precipitation measurements; they rather project in the future the data assimilation analysis of the present status of the atmosphere, which allows rainfall forecasting with a lead time of up to a few days. The accuracy of the predicted values depends both on the quality of the model and on the lead time length, due to possible divergence owing to the chaotic nature of the atmosphere.

Depending on the available measurement and modelling approaches and the nature of application, hydrologists tend to replace the unavailable areal rainfall observations, at the required space-time scales, with suitable surrogates based on *interpolation* or *downscaling* techniques. In general, interpolation is applied when ground-based rain gauge and/ or radar networks are available, while downscaling is applied when using indirect measurements from other remote sensing devices, or predicted values from atmospheric models, all of which are usually available at much coarser scales than those required in most hydrological applications. However, a sharp distinction cannot be made since interpolation and downscaling can both be incorporated in one single approach, e.g. in order to exploit jointly the information content of both remotely sensed and rain gauge data (Fiorucci *et al.*, 2001; Todini, 2001).

Interpolation and downscaling approaches have been categorised in many ways depending on the author's background and the specific field of application (see e.g. Wilby and Wigley, 1997; Xu, 1998; Kang and Ramírez, 2001). In the present context, the underlying assumed categories relate to *stochastic* (and *statistical*) and *physically based* schemes. The papers included in this volume pertain mostly to the first category, which is traditionally the most familiar within the community of hydrologists.

The physically-based approach (also referred to as *dynamical* downscaling) involves atmospheric models at the local scale that are driven by boundary conditions derived from the output of larger-scale global circulation models. All these models are widely investigated within the meteorological community and the reader is referred to the related literature for further details.

Stochastic interpolation techniques

Stochastic interpolation attempts to obtain unbiased, minimum variance estimates of precipitation at points where measurements are not available, as a function of measurements available at a number of gauged sites. Stochastic interpolation has also become an important tool for the estimation of areal rainfall, due to the linearity of the process of aggregation from point rainfall rates to areal volumes.

The early work on stochastic interpolation is due to Gandin (1970), who developed the "objective analysis" of meteorological fields. Objective analysis defines an unbiased, minimum variance estimator, based on the assumption that the covariance structure of the variable to be interpolated is known. The major advantage of the new approach is the possibility of quantifying the uncertainty in the resulting precipitation surface or precipitation volume as a function of the number and position of gauges and of the assumed spatial covariance structure. The method of Gandin has been used extensively in meteorology for data assimilation, while receiving scant attention in hydrological applications, more traditionally based on the isohyets or the Thiessen polygons methods.

In the early seventies, Matheron (1971) formalised in more appropriate mathematical terms the early work that Krige, a South African engineer, had introduced in the field of mining engineering. These works led to the development of Kriging, the optimal linear interpolation technique derived by Krige, and of the field of geostatistics, a new branch of statistical theory. Describing geostatistics falls beyond the scope of this introductory paper; however, a few comments are necessary to clarify the possible uses of Kriging for the interpolation of rainfall fields.

Similar to Gandin's objective analysis, Kriging is a linear interpolator characterised by the property of being unbiased and of minimum variance. Unfortunately, this is true conditionally upon the *a priori* knowledge of the spatial covariance (in the case of stationary random functions) or upon the knowledge of a special function, the variogram (in the case of non-stationary fields). Originally, this property was taken for granted and no investigation was made of the effect of the uncertainty induced by the estimation of such functions. Later, Kitanidis (1986) introduced a Bayesian approach to Kriging, more specifically to a special formulation of Kriging based upon the generalised covariance functions (Kitanidis, 1983), and appropriately discussed the problem of uncertainty induced by parameter estimation. Additional Bayesian approaches have been developed, among others, by Omre and Halvorsen (1989), Le and Zidek (1992) and Woodbury and Ulrych (2000). Bayesian approaches have the advantage of assessing and reducing the effects of uncertainty on model parameters, although at the expense of extensive numerical integration, generally based upon Monte Carlo or Markov chain, Monte Carlo techniques.

As an alternative to full Bayesian approaches, in view of these difficulties, a new technique is presented by Todini (2001) and demonstrated for the interpolation of the rainfall field in a region (Todini *et al.*, 2001).

Unfortunately, although Kriging is a powerful and effective tool, it does not address all the rainfall interpolation problems. As a matter of fact, observed rainfall generally shows strong anisotropy in space and, in particular, along the third co-ordinate, the elevation above sea level, where the behaviour is non-linear. In addition, the typical vertical scale (hundreds or thousands of metres) hardly compares with the tens or hundreds of kilometres of the horizontal scale. Improvements in the interpolation of rainfall will probably stem from the use of space-time scale-dependent non-linear techniques similar to the ones used in downscaling.

Stochastic models for rainfall downscaling

Stochastic downscaling aims at the reconstruction of (possible scenarios of) the small scale structure of rainfall in either the spatial or temporal domain (or both) by assuming rainfall can be suitably interpreted as a random process with specified conditions applied to some statistics that are relevant to the application of concern. Reconstruction of rainfall scenarios is therefore based on appropriate constraints. The trivial (and traditional) condition imposed on the downscaling process is that the accumulated values at some larger scale are preserved. In the paper by Fiorucci *et al.* (2001) a new approach is proposed where the above condition is relaxed, assuming that even the large scale values (when obtained as rainfall estimates, e.g. from remote sensing) can be interpreted as expected values with an associated probability distribution.

`The usual objective of stochastic downscaling is however to preserve pre-specified conditions imposed on statistical parameters related to first and second order moments of the random process, including correlation. More sophisticated models aim at preserving additional characteristics of the rain process that are relevant to hydrological applications, such as the scaling of moments and/or intermittency, clustering, fractional coverage, etc.

Cascade-based models, such as the one proposed by Güntner *et al.* (2001) for temporal downscaling of daily to hourly rainfall, became quite popular in recent times and derive from the very immediate operation of disaggregating a single large scale value into a series of lower scale contributions by partitioning and weighting the original figure while preserving the desired characteristics of the process. The main feature reproduced here is the empirically observed scaling behaviour of rainfall in both space and time. The cascade process essentially "repeatedly divides the available space (of any dimension) into smaller regions, in each step re-distributing some associated quantity according to rules specified by the so-called cascade generator" (Güntner *et al.*, 2001). The dependence of the model's parameters and the generator itself on various possible characteristics of the rainfall climate at the site of concern allows linkage with appropriate climatological features, which ensure the desired statistics are preserved and reproduced. The scale-invariant assumptions of these models are fulfilled in many climates and different model parameters allow reflecting the dominance of e.g. convective and/or advective processes in various cases.

A wide range general assessment of the relative merits and performances of cascade-based rainfall downscaling models in different geographical regions and with various governing rainfall processes is clearly required, and the paper by Güntner *et al.* (2001) is a valuable step forward in this direction. Because cascade-based models (and scaling models in general) have been supported by the need to model extreme rainfall values better, the discussion presented by the authors in their conclusions about the overestimation of extreme rainfall for frontal-dominated temperate regions deserves some attention and highlights the need for additional case studies, possibly in the space-time domain as well.

Other stochastic models for rainfall downscaling rely on various approaches based on rain field reconstruction techniques that involve direct generation of random fields with specified characteristics. Since fractional coverage of real rain fields is a relevant feature with practical implications in many hydrological studies, models have been proposed that address the two processes of rain/no-rain definition separately and that of rainfall intensity within the rainy areas.

Mackay *et al.* (2001) present a location-based approach that allows for the generation of an ensemble of disaggregated fields at each time-step for a time-series of coarse scale rainfall. The method considers the generation of the wet areas and the simulation of rainfall intensities separately. For the first task, a nearest-neighbour Markov scheme based upon a Bayesian technique used in image processing is implemented so as to preserve the structural features of the observed rainfall. The second task is dealt with by seeking to reproduce the morphological characteristics of the field of rainfall intensities through a random sampling of intensities according to a Beta distribution and their allocation to pixels chosen so that the higher intensities are more likely to be further from the dry areas. Notwithstanding some calibration and parameter assessment issues, this family of methods seems to be quite promising for practical applications.

With the increasing availability of operational, remotely-sensed estimates of space–time rainfall over wide areas, such as from RADAR and/or satellite sensors, downscaling methods that are able to include such information as a dominant constraint are becoming important for hydrological applications. Pegram and Clothier (2001) propose the String of Beads model as a space-time model of rain fields that are measured by weather RADAR. The model is driven by two auto-regressive time series models, the first one at the image scale and the other at the pixel scale, so as to give it a causal nature.

The problem of conditioning to a set of point measurements, such as rain-gauge measurements, or a random field of rainfall obtained on a lattice from meteorological RADAR or satellite image, is addressed by Fiorucci *et al.* (2001) and Todini (2001) using similar methods. These methods are based on the assumption that remotely sensed estimates of instantaneous rain fields can be viewed as probabilistic indicators of the actual rain rates and used to infer at least one parameter of the probability density function for rain intensity at each pixel site. This represents the underlying assumption of the geostatistical methodology proposed by Fiorucci *et al.* (2001), aimed at suitable integration of rainfall estimates from remote sensing with the observed figures from a number of rain gauges at the ground. Downscaling is performed by assuming the *a posteriori* estimates of the rain values at each grid cell as the *a priori* large scale conditioning values for reconstruction of the rain field at finer scale. In the paper by Todini (this issue) the point rain-gauge measurements are first extended to the pixel scale by means of Block Kriging, and successively combined with the RADAR rainfall estimates by means of a Bayesian approach, in order to provide unbiased and minimum variance *a posteriori* estimates.

Statistical downscaling methods are based on quite a different approach with respect to those mentioned above. In this case, with reference to the output of atmospheric circulation models at the synoptic or mesoscale, temporal and spatial scale details of the variables of interest at sub-grid scales — including rainfall — are derived using statistical climate inversion techniques based on regression methods (linear or non-linear). The underlying hypothesis is that meaningful statistical relationships between atmospheric state and space-time rainfall can be established at the desired scales. One of the main limitations of statistical methods is that their applicability is strictly limited to the scales at which regression relationships are obtained and to the climatic region where the regression data have been recorded. A series of case studies is therefore included in this volume with the aim of providing the reader with a view of possible approaches used in different contexts.

In the paper by Bertacchi Uvo *et al.* (2001) linear regression models based on singular value decomposition are presented with the aim of statistically downscaling atmospheric variables to estimate average rainfall on a 12-hour basis. High correlations between observed and estimated precipitation are obtained in the case study analysed by considering precipitable water and wind speeds at 850 hPa.

The variance of actual rain fields is seldom captured fully by statistical downscaling models and the small scale rainfall obtained fails to reproduce low frequency values. Despite the many explanations provided and the related solutions proposed, probably enhanced capabilities will be obtained only at the cost of developing much more complex methods, which frustrates the original valuable simplicity of the method. To mediate between the two requirements above, covariances between summer rainfall and the eigenvectors (EOFs) of summer North Atlantic sea surface temperature (SST) anomalies are used by Wilby (2001) so as to evaluate the skill of summer rainfall forecasts produced by statistical downscaling models that use EOFs of the preceding winter SSTs.

Empirical Orthogonal Functions (EOFs) is the method employed by Rodriguez-Puebla *et al.* (2001) to describe the variance distribution and to summarise precipitation data over the Iberian Peninsula into a few modes. The time series associated with these modes are then analysed with large scale circulation indices and tropical sea surface temperature anomalies by using lag cross-correlation and cross-spectrum.

Conclusions

This volume collects together a series of papers on rainfall interpolation and downscaling techniques that has the objective of providing an overview of recent developments in the field. One of the expectations of the guest editors at the beginning of this effort was that of highlighting some current issues and possible hints for future improvements of the available models as well as for better understanding of the rainfall process. Therefore, the objective of this introductory paper is to provide the reader with a logical thread connecting all the papers included in this volume and the editors' interpretation of and a few considerations on the main issues treated.

A common problem of all interpolation and downscaling methods remains model validation. Despite the many case studies presented, direct validation of downscaling models is strictly impossible (except for methods focused only on the rainfall process in time) due to the lack of reliable measurements of areal rainfall and the related patterns in both space and time. Indirect validation is therefore widely resorted to and hydrological rainfall-runoff models are sometimes used to transform the variability of rainfall in space and time into simulated flow discharges in river courses, the latter being suitably measured by means of any flow gauge. Obviously, this introduces additional uncertainties related to the many processes involved in the rainfall-runoff transformation and the modelling scheme used, and makes the validation of downscaling models only affordable in an indirect and therefore scarcely reliable way. To overcome this difficulty, standard synthetic tests based on Monte Carlo approaches should be set up and, most important, agreed upon by the hydrological and meteorological communities, to test objectively the different approaches in terms of expectation and reduction in uncertainty (Todini, 2001). Nonetheless, the investigation of detailed case studies is of great relevance and must be encouraged. In particular, case studies involving more than one method for rainfall downscaling are necessary and detailed inter-comparisons required. Experiments should be organised and the performances of the various methodologies available be compared against both the synthetic and the indirect validation technique. For practical applications, indeed, it would be useful to compare the computational burden and complexity of the best performing methods against their capabilities in reproducing the relevant features of interest related to the space-time variability of rainfall.

References

Bertacchi Uvo, C., Olsson, J., Morita, O., Jinno, K., Kawamura, A., Nishiyama, K., Koreeda, N. and Nakashima, T., 2001. Statistical atmospheric downscaling for rainfall estimation in Kyushu Island, Japan, *Hydrol. Earth. System Sci.*, **5**,259–271.

Fiorucci, P., La Barbera, P., Lanza, L.G. and Minciardi, R., 2001. A geostatistical approach to multisensor rain field reconstruction and downscaling, *Hydrol. Earth. System Sci.*, **5**, 201–213.

Humphrey, M.D., Istok, J.D., Lee, J.Y., Hevesi, J.A., and Flint, A., 1997. A new method for automated dynamic calibration of tipping-bucket rain gauges, *J. Atmos. Ocean. Technol.*, **14**, 1513–1519.

Gandin, L.S., 1970. *The Planning of Meteorological Station Networks*, (Technical Note No. 111). WMO No. 265, Geneva

Güntner, A., Olsson, J., Calver, A. and Gannon, B., 2001. Cascade-based disaggregation of continuous rainfall time series: the influence of climate, *Hydrol. Earth. System Sci.*, **5**, 145–164.

Kang, B. and Ramírez, J.A., 2001. Comparative study of the statistical features of random cascade models for spatial rainfall downscaling, In*: Proc. AGU Hydrology Days 2001*, J.A. Ramírez (Ed.), Hydrology Days Publications, Fort Collins, CO., 151–164.

Kitanidis, P.K., 1983. Statistical estimation of polynomial generalised covariance functions and hydrologic applications, *Water Resour. Res*., **19,** 909–921.

Kitanidis, P.K., 1986. Parameter uncertainty in estimation of spatial functions: Bayesian analysis, *Water Resour. Res.*, **22**, 499–507.

La Barbera, P., Lanza, L.G. and Stagi, L., 2001. Tipping bucket mechanical errors and their influence on rainfall statistics and extremes, *Water Sci. Technol.*, submitted.

Le, N.D. and Zidek, J.V., 1992. Interpolation with uncertain spatial covariances: a Bayesian alternative to Kriging, *J. Multivariate Anal.*, **43**, 351–374.

Linsley, R.K., Kholer, M.A. and Paulhus, J.L.H., 1949. *Applied Hydrology*, McGraw Hill, New York.

Mackay, N.G., Chandler, R.E., Onof, C. and Wheater, H.S., 2001. Disaggregation of spatial rainfall fields for hydrological modelling, *Hydrol. Earth. System Sci.*, **5**, 165–173.

Matheron, G., 1971. *The theory of regionalised variables and its applications*, Cahiers du Centre de Morphologie Mathematique, No. 5, Fontainebleau, France.

Matheron, G., 1981. Splines and Kriging: their formal equivalence, In: *Down-to Earth Statistics: Solutions Looking for Geological Problems*, D.F.Merriam (Ed.), Syracuse University Geological Contributions, Syracuse. 77–95.

Omre, H. and Halvorsen, K.B., 1989. The Bayesian bridge between simple and universal Kriging, *Math. Geol.*, **21**, 767–786.

Pegram, G.S. and Clothier, A.N., 2001. Downscaling rainfields in space and time, using the String of Beads model in causal mode, *Hydrol. Earth. System Sci.*, **5**, 175–186.

Rainbird, A.F., 1967. *Methods of estimating areal average precipitation*, Reports on WMO/IHD Projects, Report no. 3.

Rodriguez-Puebla, C., Encinas, A.H. and Nieto, S., 2001. Seasonal precipitation prediction over the Iberian peninsula based on statistical modelling, *Hydrol. Earth. System Sci.*, **5**, 233–244.

Sevruk, B., 1982. Methods of correction for systematic error in point precipitation measurement for operational use, *Operational Hydrology Rep. No. 21*, WMO Rep. No. 589.

Todini, E., 2001a. A Bayesian technique for conditioning radar precipitation estimates to rain gauge measurements, *Hydrol. Earth. System Sci.*, **5**, 187–199.

Todini, E., 2001b. Influence of parameter estimation uncertainty in Kriging. Part 1 – Theoretical Development, *Hydrol. Earth. System Sci.*, **5**, 215–223.

Todini, E., Pellegrini, F. and Mazzetti, C., 2001. Influence of parameter estimation uncertainty in Kriging. Part 2 – Test and case study applications, *Hydrol. Earth. System Sci.*, **5**, 225–232.

Xu, C-Y., 1998. *From GCMs to river flow: A review of downscaling methods and hydrologic modelling approaches*, Uppsala University, Rep. Series A No. 54.

Wilby, R.L., 2001. Downscaling summer rainfall in the U.K. from north Atlantic ocean temperature, *Hydrol. Earth. System Sci.*, **5**, 245–257.

Wilby, R.L. and Wigley, T.M.L. 1997. Downscaling general circulation model output: a review of methods and limitations, *Prog. Phys. Geog.*, **21**, 530–548.

WMO, 1986. *Manual for Estimation of Probable Maximum Precipitation*, Operational Hydrology Report n.1, WMO N. 332, Geneva.

WMO, 1994. *Guide to Hydrological Practices*. WMO N. 168, Geneva.

Woodbury, A.D. and Ulrych, T.J., 2000. A full-Bayesian approach to the groundwater inverse problem for steady state flow, *Water Resour. Res.*, **36**, 2081–2093.